Wide Area Information Servers:
Improving Access to Information

By Eliot J. Christian and Timothy L. Gauslin

In many instances, scientists, policy makers, and educators as well as the general public would make better use of computerized information if it were easier to access. What can be done to simplify access to data and information?

In the late 1980's, the library and information services community reached consensus on an access standard. This open standard, known in the United States as the Information Retrieval Service Definition and Protocol Specification (Z39.50), defines a uniform method for access to data and information. The idea is that an open standard permits data and information to be stored on computer systems having very different hardware and software, yet a user can retrieve information from any computer in the same way.

The USGS is working with a consortium including several major corporations and about 150 universities to develop a version of Z39.50 that is available as public domain software, known as Wide Area Information Servers (WAIS). WAIS is targeted to users who have limited computer skills and need a powerful search and retrieval mechanism. This approach is already being used extensively in the USGS and is gaining wide acceptance in the Federal Government as a whole.

In a typical WAIS search, the user selects information servers or searches directories of servers, using familiar English words. That search is formatted to the Z39.50 standard, and the information servers match the search words to the server contents. The user sees a list of titles found and then views those of interest (Figure 1).

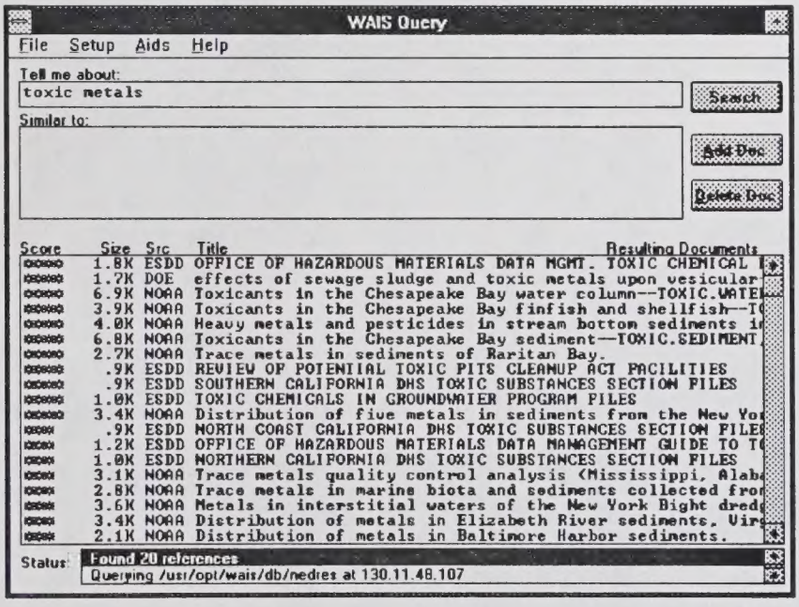
WAIS Query
File Setup Aids Help
Tell me about:
toxic metals
Search
Similar to:
Add Doc
Delete Doc
Score Size Src Title Resulting Documents
1.8K ESDD OFFICE OF HAZARDOUS MATERIALS DATA MGMT. TOXIC CHEMICAL
1.7K DOE effects of sewage sludge and toxic metals upon vesicular
6.9K NOAA Toxicants in the Chesapeake Bay water column--TOXIC.WATE
3.9K NOAA Toxicants in the Chesapeake Bay finfish and shellfish--T
4.0K NOAA Heavy metals and pesticides in stream bottom sediments i
6.8K NOAA Toxicants in the Chesapeake Bay sediment--TOXIC.SEDIMENT
2.7K NOAA Trace metals in sediments of Raritan Bay.
.9K ESDD REVIEW OF POTENTIAL TOXIC PITS CLEANUP ACT FACILITIES
.9K ESDD SOUTHERN CALIFORNIA DHS TOXIC SUBSTANCES SECTION FILES
1.0K ESDD TOXIC CHEMICALS IN GROUNDWATER PROGRAM FILES
3.4K NOAA Distribution of five metals in sediments from the New Yo
.9K ESDD NORTH COAST CALIFORNIA DHS TOXIC SUBSTANCES SECTION FILE
1.2K ESDD OFFICE OF HAZARDOUS MATERIALS DATA MANAGEMENT GUIDE TO T
1.0K ESDD NORTHERN CALIFORNIA DHS TOXIC SUBSTANCES SECTION FILES
3.1K NOAA Trace metals quality control analysis (Mississippi, Alab
2.8K NOAA Trace metals in marine biota and sediments collected fro
3.6K NOAA Metals in interstitial waters of the New York Bight dred
3.4K NOAA Distribution of metals in Elizabeth River sediments, Vir
2.1K NOAA Distribution of metals in Baltimore Harbor sediments.
Status: Found 20 references
Querying /usr/opt/wais/db/nedres at 130.11.48.107

Figure 1. Multiple sources can be selected for searching, regardless of their different formats.

Since users do not know which words would be key to finding everything that may be relevant, WAIS provides an approach known as "relevance feedback." A user simply highlights any retrieved text that seems especially relevant and WAIS uses that piece of text to do further searching.

The USGS began its involvement with WAIS through an initiative to enhance the Earth Science Data Directory (ESDD). ESDD is a useful source for references to earth science data, including many at the State level, and is also a comprehensive list of data holdings relevant to Arctic research. WAIS is

especially appropriate for ESDD because the ESDD user community ranges from local citizens to international global change researchers. The ability of WAIS to place the ESDD in the context of other information sources is especially powerful for these users. WAIS is also expected to be used by the USGS and others for directories of spatial data and for information related to water and geologic data.

The USGS has added features to WAIS that are useful for ESDD users (for example, location searching, phrase searching, and key-word searching within fields) in a manner compliant with the Z39.50 standard and in concert with the network of WAIS developers. In fact, the WAIS software has many examples of contributed parts from hundreds of developers worldwide. By maintaining compliance with the standard, users of the USGS/WAIS software can access any Z39.50 server but have additional capabilities when accessing the USGS server. Conversely, the USGS information servers are accessible to all users of Z39.50, whether they received software from USGS or not.

For the location searching feature, the USGS developed the software to map the latitude/longitude referenced in the text and to allow a user to search by drawing a polygon on a map (Figure 2). The University of North Carolina developed the server software that accepts the polygon vertices and matches them to source entries.

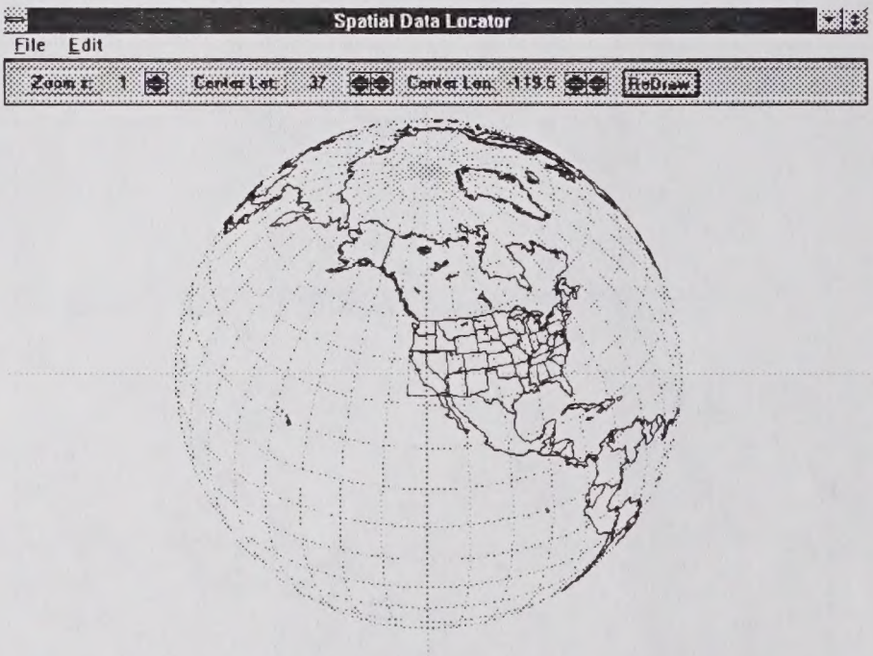

Figure 2. A USGS enhancement to WAIS provides for locating data spatially.

As the Z39.50 standard becomes widely used, opportunities arise to simplify interaction with data and information repositories. For example, Apple Computers created a dynamic searching mechanism that periodically retrieves requested information automatically and prepares a "personalized newspaper" showing updated information that the user requested. The USGS has also added a simple way for a user to suspend the WAIS session and access an existing data system such as the USGS Global Land Information System. Another idea is to give the WAIS user an option to contribute data and information products on-line. WAIS would help the contributor to document the product appropriately and would then route the document and product electronically through peer review.

Versions of WAIS are already available for most computers, and there are hundreds of information servers freely accessible. WAIS and Z39.50 are being used not only for bibliographic information but also for providing on-line access to data. As an integrating technology, WAIS promises to vastly simplify the problems of access to scientific and technical data and information.